Leidos-SERAS
2890 Woodbridge Avenue, Building 209 Annex
Edison, NJ 08837-3679
Telephone: 732-321-4200, Facsimile: 732-494-4021

DATE: November 29, 2018

TO: Felicia Barnett, Director SCMTSC, EPA Work Assignment Manager

FROM: Donna Getty, SERAS Statistician

THROUGH: Richard Leuser, SERAS Deputy Program Manager/Task Leader

SUBJECT: STATISTICAL COMMENTS ON DRAFT FINAL UNITED STATES NAVY PARCEL G REMOVAL SITE EVALUATION WORK PLAN ISSUED NOVEMBER 2018, HUNTERS POINT NAVAL SHIPYARD SITE (HPNS), SERAS-106, WORK ORDER #83

INTRODUCTION

United States (US) Environmental Protection Agency (EPA) Region 9 personnel requested a final statistical review of the sampling strategies outlined in the US Navy's *Draft Final Parcel G Removal Site Evaluation Work Plan Former Hunters Point Naval Shipyard, San Francisco, California* (November 2018) (Work Plan). This version of the Naval Work Plan was expected to address comments made by the US EPA Regulators, the City of San Francisco, and the general public on the August 2018 version of the Draft Work Plan. Both the June 2018 and the current November 2018 versions of the work plan present an overview of proposed methodologies and protocols for the investigation of radiological contamination in Parcel G and strategies for establishing background radiological levels, building on the strategies proposed in the over-arching Draft Work Plan put forth by the Navy in February 2018.

The proposed work is being conducted with the current knowledge of allegations that previous sampling and analysis efforts conducted by a Navy subcontractor may have included sample collection fraud, data manipulation, improper sample custody, and other issues that indicate that radiological contamination may still be present on Parcel G. Comments contained in this review refer back to statistical reviews provided by Leidos on earlier versions of the Work Plan. For additional statistical details of the proposed sampling strategies refer to the Leidos Technical Memoranda (TM) delivered to the US EPA in February, March, and August 2018.

DISCUSSION

Review of the November 2018 version of the Work Plan for Parcel G demonstrates that the majority of comments provided in the Leidos August 2018 TM on the June 2018 version of the Work Plan were addressed, however one critical issue was not corrected/addressed. In the Navy's over-arching Work Plan

provided to the EPA Regulators in February 2018, the Type I error, represented as α, and Type II error, represented as β were set at 0.01 (99% confidence) for the statistical testing. The EPA Regulators concurred with these proposed levels. These statistical levels were then used by the EPA Regulators in the calculation of the minimum number of samples, to be collected on a systematic sampling grid per Trench Unit (TU) or Sampling Unit (SU). This required number of samples to be collected would have a 1% chance of declaring whether a TU or SU meets compliance when in reality it does not (false negative) and only a 1% chance of declaring whether a TU or SU does not meet compliance when it actually does (false positive). These levels have also been a driving factor in the design of and the agreement to the sampling strategies the US EPA and Navy have been discussing since February 2018. Discussions of sample size between the US EPA and the Navy have repeatedly addressed the need for a minimum of 25 samples per TU or SU.

Confusing the issue is that sample sizes computed thus far have been based on the use of a Wilcoxon Rank Sum (WRS) test to compare the Reference Background Area (RBA) data to the individual TU/SUs to determine compliance; when to date, the means for establishing compliance has not been established or agreed to. These calculations have been a means to develop a starting sampling size until new data can be collected to refine the design. It should also be noted, that the current sampling design calls for 25 surface samples to be collected per RBA and the results from the RBA are expected to be less variable than the on-site results.

On pages 3-6 and 4-5 of the recent November 2018 Navy Draft Final Work Plan, the α-level remains set at 0.01, however the β-level has been altered to 0.05. The β-level of 0.05 is used by the Navy to justify an initial minimum sample size of 18 samples for the TU/SU's designated as MARSSIM Class 1 units in the Work Plan. MARSSIM refers to the sampling strategies and protocols detailed in the *Multi-Agency Radiation Survey and Site Investigation Manual* (MARSSIM, 2002). The Work Plan references MARSSIM as the source for obtaining a sample size of 18 and documents the calculation and assumptions made. However, had the β-level been set to the previously agreed upon 0.01, these calculations would result in a sample size of 25.

The justification of 25 as the required sample size has been discussed in previous Leidos' TMs. Below is text provided in the August 2, 2018 Leidos TM. Page numbers refer to the June 2018 version of the Navy Work Plan for Parcel G.

> *"Section 3.4.1, page 3-4 states that the number of samples to be collected per TU/ SU will follow the previously established protocol of a minimum of 18 systematically located samples. A discussion of sample size was presented by Leidos on March 15, 2018 demonstrating that following the historical protocol of n=18 samples may not achieve the desired confidence level and associated power of 99% (α=0.01 and β=0.01)."*
>
> *"Calculations presented in the Leidos March 2015 TM demonstrate that n=18 may be an insufficient number of samples for conducting a WRS Test with 99% confidence. In the TM, samples sizes were computed based on the standard deviations computed from the historical sampling results of the original RBA's. Standard deviations ranged from 0.27 to 0.47 picoCuries per gram (pCi/g) for Radium-226 and from 0.031 to 0.046 pCi/g for Cesium-137 in the historical RBAs. Expecting greater variability in the on-site data than in the RBA data, and to be conservative and protective of human health, the maximum standard deviations from the RBAs were used in calculating sample size for α=0.01, and β=0.01. Applying the MARSSIM methodology a sample size of n=25 was computed."*

CONCLUSIONS

Sample size calculations for statistical comparisons are based on several factors: error levels (α and β), the known or projected standard deviation of the data, and the data quality objectives (DQOs) which include the type of statistical tests which will be conducted. Sample sizes in the Navy's November 2018 Work Plan are based on the use of a Wilcoxon Rank Sum (WRS) test with α=0.01 and β=0.05. To date, the US EPA has not agreed to the use of the WRS test but this reviewer has recognized the sample size calculations as a "jumping off point" to permit sampling to begin. It is this reviewer's understanding that the US EPA has requested a point-to-actionable level comparison to determine compliance. This could be achieved through a comparison of individual point data to a statistically computed background threshold value (BTV), such as an upper tolerance limit (UTL) or upper confidence limit (UCL). It is recommended that ProUCL software is used to identify the most representative BTV to use once new sampling results are obtained. The initial sampling results should undergo statistical examination including the identification of potential outliers, identification of the most likely population distribution and computations of skewness and variability. It is necessary to have a large enough sample size at the beginning of the sampling to establish these parameters and therefore the required sample size for each TU and SU that will achieve the agreed upon DQOs.

Note: Calculations based on the historical data demonstrate that 18 samples will only provide approximately 60% confidence that 95% of the unsampled soil in a TU meets compliance if a UTL is chosen as the BTV.

Based on the statistical calculations documented in previous Leidos TMs, the DQOs for HPNS and the previously agreed upon error levels of α=0.01 and β=0.01, the initial recommended minimum sample size for soils is still n=25, for TUs or SUs identified for Class1 MARSSIM-based sampling on Parcel G at HPNS. As discussed with the Navy, the minimum required sample size may change following the sampling of the RBAs which are planned to occur prior to on-site sampling or following the sampling of the initial TUs/SUs if conducted simultaneously with the RBA sampling. If standard deviations of the analytical results for Radium-226 prove to be greater than the historical sampling results from which n=25 was computed, minimum sample size may increase. If the standard deviations are less than the historical sampling results, sample size may decrease. However, the agreed upon α- and β-levels will not be affected by the RBA sampling events. Both should continue to be set at 0.01, as previously agreed upon, to facilitate the determination of how the RBA data will be compared to the on-site individual TU/SU data.

REFERENCES

Leidos. Technical Memorandum. *Final—Statistical Review of United States Navy Proposed Work Plan for Radiological Survey and Sampling, Hunters Point Naval Shipyard Site (HPNS)*. Work Assignment SERAS-106, Work Order #83. March 15, 2018

Leidos. Technical Memorandum. *Final - Statistical Review of United States Navy Draft Parcel G Removal Site Evaluation Work Plan,* Hunters Point Naval Shipyard Site (HPNS), San Francisco, CA. Work Assignment SERAS-106, Work Order #83. August 2, 2018.

Leidos. Technical Memorandum. *Final—Statistical Review of United States Navy Proposed Work Plan for Radiological Survey and Sampling.* Hunters Point Naval Shipyard Site (HPNS), San Francisco, CA. Work Assignment SERAS-106, Work Order #83. March 15, 2018.

Leidos. Technical Memorandum. *Exploration of a Potential Method for Establishing a Level of Confidence across a Parcel that Radioactivity Levels Are Below Actionable Levels.* Hunters Point Naval Shipyard Site (HPNS), San Francisco, CA. Work Assignment SERAS-106, Work Order #83. February 22, 2018.

MARSSIM. *Multi-Agency Radiation Survey and Site Investigation Manual (Revision 1).* NUREG-1575 Rev. 1, EPA 402-R-97-016 Rev. 1, DOE EH-0624 Rev. 1, August 2002. (accessed March 12, 2018). https://www.epa.gov/radiation/multi-agency-radiation-survey-and-site-investigation-manual-marssim

Naval Facilities Engineering Command (NAVFAC). *Draft Parcel G Removal Site Evaluation Work Plan, Former Hunters Point Naval Shipyard.* San Francisco, CA. June 2018.

Naval Facilities Engineering Command (NAVFAC). *Draft Work Plan, Radiological Survey and Sampling, Former Hunters Point Naval Shipyard.* San Francisco, CA. February 2018.

Nuclear Regulatory Commission (NRC). *Nonparametric Statistical Methodology for the Design and Analysis of Final Status Decommissioning Surveys.* Interim Draft Report for Comment and Use. U.S. Nuclear Regulatory Commission Office of Nuclear Regulatory Research. NUREG-1505, Rev. 1. 1998. https://www.nrc.gov/docs/ML0618/ML061870462.pdf (accessed on 2/18/2018)

U.S. Environmental Protection Agency (EPA). ProUCL 5.1 User Guide. Statistical Software for Environmental Applications for Data Sets with and without Nondetect Observations. EPA/600/R-07/041. 2015. https://www.epa.gov/land-research/proucl-version-5100-documentation-downloads

cc: Central File - WA # SERAS-106
Electronic File - I:/Archive/SERAS/106/D/TM/112918
Paul Carter, SERAS Program Manager (cover page only)